

ENTROPIA E DISTRIBUIÇÕES CONJUGADAS

António G. Portela (*)

Bento J. F. Murteira (**)

1 — Entropia

O conceito de entropia, estudado normalmente no âmbito da teoria da informação, introduz-se de forma conveniente a partir das distribuições finitas. Dado um conjunto finito de números reais, $\Theta = \{\theta_i: i = 1, 2, \dots, n\}$, seja $h(\theta)$ uma distribuição finita (função de probabilidade) sobre Θ ; a entropia de h , simbolicamente, $\mathcal{E}[h]$, é definida por:

$$(1.1) \quad \mathcal{E}[h] = - \sum_{i=1}^n h(\theta_i) \log h(\theta_i)$$

onde \log significa logaritmo neperiano e se toma $h(\theta_i) \cdot \log h(\theta_i) = 0$ sempre que seja $h(\theta_i) = 0$.

A entropia mede a quantidade de incerteza inerente à distribuição h . Para compreender esta asserção, considerem-se dois casos extremos, um em que a incerteza é mínima, outro em que é máxima. A incerteza é mínima quando se tem $h(\theta) \equiv h_m(\theta)$, com:

$$(1.2) \quad h_m(\theta_i) = 1, h_m(\theta_j) = 0, j = 1, 2, \dots, n \dots j \neq i$$

De facto, com a distribuição singular h_m , tem-se a certeza de que ocorre θ_i e de que não ocorre nenhum dos outros $\theta_j, j \neq i$. Sai então:

$$(1.3) \quad \mathcal{E}[h_m] = 0$$

valor mínimo da entropia. A incerteza é máxima quando se tem $h(\theta) \equiv h_M(\theta)$, com:

$$(1.4) \quad h_M(\theta_i) = 1/n, i = 1, 2, \dots, n$$

De facto, com a distribuição uniforme h_M , a situação é extremamente vaga ou difusa, pois os θ_i ocorrem com igual probabilidade. Sai então:

$$(1.5) \quad \mathcal{E}[h_M] = - \sum_{i=1}^n (1/n) \log (1/n) = \log n$$

que se demonstra (Ventzel, 1973) ser o valor máximo da entropia.

A extensão do conceito de entropia às distribuições contínuas envolve algumas dificuldades (Pugachev, 1973, e Ventzel, 1973). Dado um conjunto contínuo, $\Theta \subset R$, seja $h(\theta)$ uma função de densidade sobre Θ ; para medir a indeterminação do fenómeno aleatório caracterizado por h , pode usar-se o conceito de entropia reduzida:

$$(1.6) \quad \mathcal{E}^*[h] = - \int_{\Theta} h(\theta) \log h(\theta) d\theta$$

(*) Professor catedrático do Instituto Superior Técnico.

(**) Professor catedrático do Instituto Superior de Economia.

desde que se tenha presente que a entropia propriamente dita é medida a partir de uma origem que depende de $\Delta \theta$, isto é, da precisão com que se observa a variável θ (Ventzel, 1973). Desde que não haja lugar para confusão, escreve-se $\mathcal{E}[\cdot]$, em vez de $\mathcal{E}^*[\cdot]$.

2 — Distribuições conjugadas

As distribuições conjugadas possuem grande interesse prático na inferência bayesiana.

Num problema de inferência bayesiana, seja Θ o espaço dos parâmetros, $h(\theta)$ uma distribuição *a priori* sobre Θ e

$$(2.1) \quad \mathcal{F} = \{f(x | \theta) : \theta \in \Theta\}, \quad x \in \mathcal{X}'$$

a família de modelos probabilísticos referentes à experiência em que se observa a variável ou vector aleatório, X , com espaço de resultados \mathcal{X}' , é função de probabilidade ou de densidade, $f(x|\theta)$. É bem sabido que, observado $X = x$, a distribuição *a posteriori* de θ , $h(\theta|x)$, dada pelo teorema de Bayes:

$$(2.2) \quad h(\theta | x) = f(x | \theta) h(\theta) / \int_{\Theta} f(x | \theta) h(\theta) d\theta$$

é instrumento básico na construção de inferências.

Uma família de distribuições:

$$(2.3) \quad \mathcal{H} = \{h(\theta; \lambda) : \lambda \in \Lambda\}, \quad \theta \in \Theta$$

onde λ é um parâmetro, escalar ou vector, diz-se conjugada de \mathcal{F} se e somente se é fechada em relação a (2.2), isto é, se e somente se:

$$(2.4) \quad h(\theta) \equiv h(\theta; \lambda_0) \in \mathcal{H} \text{ implica } h(\theta | x) \equiv h(\theta; \lambda_1) \in \mathcal{H}$$

onde, consequentemente, $\lambda_0 \in \Lambda$ e $\lambda_1 \in \Lambda$.

Caso importante é aquele em que θ é uma probabilidade, $\Theta = [0,1]$, e a experiência consiste na realização de um número fixo, N , de provas de Bernoulli; a família de modelos probabilísticos, a binomial:

$$(2.5) \quad \mathcal{F}_B = \{f(x | \theta) = \binom{N}{x} \theta^x (1 - \theta)^{N-x} : N \text{ int.}^\circ \text{ pos.}, 0 \leq \theta \leq 1\}, \quad \mathcal{X}' = \{0, 1, \dots, N\}$$

tem por conjugada a família das distribuições beta:

$$(2.6) \quad \mathcal{H}_B = \{h(\theta; \alpha, \beta) = \theta^{\alpha-1} (1 - \theta)^{\beta-1} / B(\alpha, \beta) : \alpha > 0, \beta > 0\}, \quad 0 \leq \theta \leq 1$$

onde:

$$B(\alpha, \beta) = \int_0^1 \theta^{\alpha-1} (1 - \theta)^{\beta-1} d\theta$$

É fácil de verificar — através da aplicação de (2.2) — que, se a distribuição *a priori* é uma beta:

$$h(\theta) \equiv h(\theta; \alpha_0, \beta_0) = \theta^{\alpha_0-1} (1 - \theta)^{\beta_0-1} / B(\alpha_0, \beta_0)$$

a distribuição *a posteriori* é também uma beta, $h(\theta|x) \equiv h(\theta; \alpha_1, \beta_1)$, com:

$$(2.7) \quad \alpha_1 = \alpha_0 + x, \beta_1 = \beta_0 + N - x$$

Em geral, quando $h(\theta) \equiv h(\theta; \lambda_0)$ e $h(\theta|x) \equiv h(\theta; \lambda_1)$, $\lambda_0, \lambda_1 \in \Lambda$, o procedimento bayesiano que opera a passagem da distribuição *a priori* para a distribuição *a posteriori* pode representar-se simbolicamente:

$$(2.8) \quad \lambda_0 \xRightarrow{\mathcal{F}} \lambda_1$$

destacando o papel representado pela experimentação, isto é, pela família \mathcal{F} . A (2.8) pode dar-se a forma alternativa:

$$(2.9) \quad \lambda_0 \xRightarrow{\mathcal{F}} \lambda_0 + (\lambda_1 - \lambda_0)$$

onde $\lambda_1 - \lambda_0$ exprime a variação sofrida por λ_0 em resultado da informação decorrente do processo de amostragem. Note-se que (2.9) permite avaliar, em certo sentido, o peso da informação *a priori* relativamente à informação experimental.

No caso da binomial, tem-se:

$$(2.10) \quad (\alpha_0, \beta_0) \xRightarrow{\mathcal{F}_B} (\alpha_1, \beta_1)$$

ou, atendendo a (2.7):

$$(2.11) \quad (\alpha_0, \beta_0) \xRightarrow{\mathcal{F}_B} (\alpha_0, \beta_0) + (x, N - x)$$

como se representa na figura 1. A observação de $X = x$ significa, neste caso, que em N provas de Bernoulli se registaram x «sucessos» e $N - x$ «insucessos». Nestes termos, pode pensar-se (Barnett, 1973) que a informação *a priori* equivale a α_0 «sucessos» e β_0 «insucessos» e que a informação *a posteriori* equivale a $\alpha_1 = \alpha_0 + x$ «sucessos» e $\beta_1 = \beta_0 + N - x$ «insucessos».

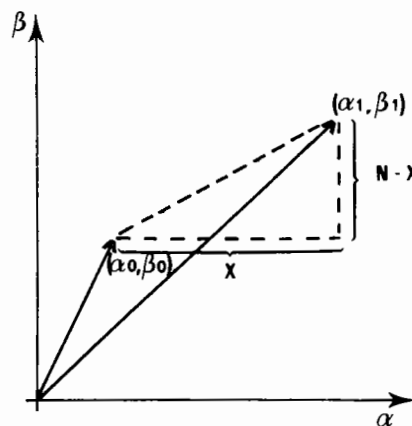


Figura 1

3 — Entropia e distribuições conjugadas

Que se passa com a entropia quando o operador bayesiano permite passar da distribuição *a priori* para a distribuição *a posteriori* com base na informação adicional fornecida pelo resultado particular da experiência?

Pode concluir-se, precipitadamente, que é sempre:

$$(3.1) \quad E[h(\theta)] \geq \mathcal{E}[h(\theta | x)]$$

argumentando que a informação experimental deve reduzir a incerteza, logo também a entropia. No entanto, a relação (3.1) não é necessariamente verdadeira, como pode facilmente verificar-se; aliás, Lindley (1956), ao notar que um resultado experimental particular pode aumentar a entropia, afirma:

This can happen when a «surprising» value of x occurs; granted the correctness of the experimental technique, the «surprising» may result in our being less sure about θ than before the experiment.

Indo mais longe, Lindley (1956) demonstra que a relação que se verifica é a seguinte:

$$(3.2) \quad \mathcal{E}[h(\theta)] \geq \int_{\mathcal{X}} \mathcal{E}[h(\theta | x)] f(x) dx$$

onde o segundo membro representa a entropia *a posteriori* média obtida sobre o espaço dos resultados.

Outro caso de muito interesse diz respeito à família normal (variância dada):

$$(3.3) \quad \mathcal{F}_N = \{f(x | \mu, \sigma) = 1/\sigma\sqrt{2\pi} e^{-(x-\mu)^2/2\sigma^2}; -\infty < \mu < +\infty, \sigma > 0\}, \mathcal{X} = (-\infty, +\infty),$$

cujas conjugadas são as próprias normais. Recorrendo mais uma vez ao teorema de Bayes, pode mostrar-se que, se $h(\mu) \equiv f(\mu_0, \sigma_0^2)$ (*) e se a experiência consiste na observação de uma amostra casual de dimensão N , $X = (X_1, X_2, \dots, X_N)$, X_i i. i. D., $X_i \equiv f(\mu, \sigma^2)$ (**), então, dada a amostra particular, $x = (x_1, x_2, \dots, x_N)$, obtém-se $h(\mu|x) \equiv f(\mu_1, \sigma_1^2)$, com:

$$(3.4) \quad \begin{aligned} \mu_1 &= [(1/\sigma_0^2)\mu_0 + (N/\sigma^2)\bar{x}] [(1/\sigma_0^2) + (N/\sigma^2)]^{-1}, \quad (\bar{x} = \sum_{i=1}^N x_i / N), \\ \sigma_1^2 &= [(1/\sigma_0^2) + (N/\sigma^2)]^{-1} \end{aligned}$$

Em esquema:

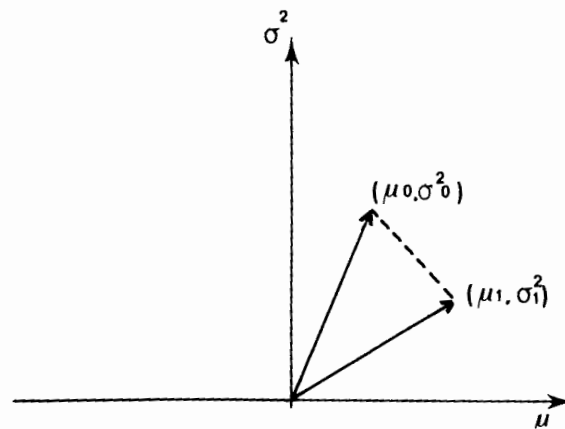
$$(\mu_0, \sigma_0^2) \xrightarrow{\mathcal{F}} (\mu_0, \sigma_0^2) + (\mu_1 - \mu_0, \sigma_1^2 - \sigma_0^2)$$

(*) Leia-se: $h(\mu)$ é uma distribuição normal com média μ_0 e desvio padrão σ_0 .

(**) Leia-se: as variáveis aleatórias, X_i , $i = 1, 2, \dots, N$, são independentes e identicamente distribuídas segundo uma normal com média μ e desvio padrão σ .

onde:

$$(3.5) \quad \begin{aligned} \mu_1 - \mu_0 &= (N/\sigma^2) (\bar{x} - \mu_0) \\ \sigma_1^2 - \sigma_0^2 &= (N\sigma_0^2/\sigma^2) [(1/\sigma_0^2) + (N/\sigma^2)]^{-1} \end{aligned}$$



CASO $\bar{x} > \mu_0$

Figura 2

O sinal de $\mu_1 - \mu_0$ é idêntico ao sinal de $\bar{x} - \mu_0$, quer dizer, $\mu_1 \geq \mu_0$, quando $\bar{x} \geq \mu_0$, mas, $\sigma_1^2 < \sigma_0^2$, sempre (v. fig. 2). A entropia da normal $I(\mu, \sigma^2)$ é de fácil cálculo (Ventzel, 1973) e tem a expressão:

$$(3.6) \quad \log(\sqrt{2\pi e \sigma^2})$$

que permite concluir que no espaço dos parâmetros:

$$\Theta = \{(\mu, \sigma^2): -\infty < \mu < +\infty, \sigma > 0\}$$

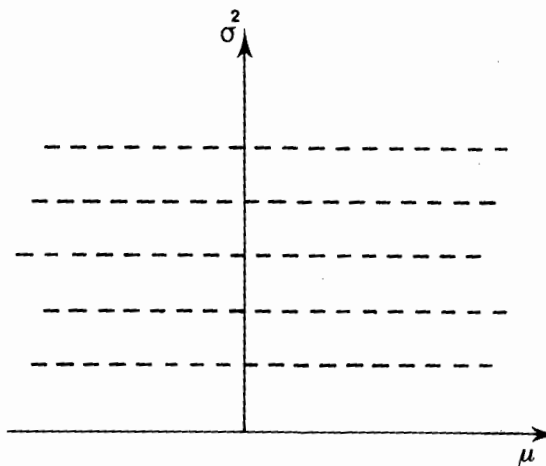


Figura 3

as isoentrópicas — linhas de igual entropia — são rectas paralelas ao eixo das abcissas (fig. 3). Dado que $\sigma_1^2 < \sigma_0^2$, há sempre redução de entropia, qualquer que seja o resultado experimental \underline{x} (comparem-se as figs. 2 e 3); mais precisamente, fixada a dimensão da amostra, N , e a variância dos X_i ($i = 1, 2, \dots, N$) — ou, o que é o mesmo, fixado N e a precisão experimental —, a redução de entropia é independente do resultado particular da experiência, porquanto:

$$\mathcal{E}[h(\mu)] - \mathcal{E}[h(\mu | x)] = \log(\sigma_0 / \sigma_1) > 0$$

Voltando ao caso da binomial, joga-se de interesse ter uma ideia da forma das isoentrópicas. A entropia da beta — família conjugada da binomial — é dada pela expressão:

(3.7)

$$\begin{aligned} \mathcal{E}[h(\theta; \alpha, \beta)] &= \int_0^1 h(\theta; \alpha, \beta) \cdot \log h(\theta; \alpha, \beta) d\theta \\ &= \log[\Gamma(\alpha + \beta) / \Gamma(\alpha) \Gamma(\beta)] + (\alpha - 1) [\Psi(\alpha) - \Psi(\alpha + \beta)] + \\ &\quad + (\beta - 1) [\Psi(\beta) - \Psi(\alpha + \beta)] \end{aligned}$$

onde:

$$\Psi(u) = d \log \Gamma(u) / du$$

Lindley (1956), utilizando as relações assintóticas:

$$(3.8) \log \Gamma(u) \sim (1/2) \log 2\pi - u + [u - (1/2)] \log u, \Psi(u) \sim \log u - (1/2u)$$

obteve a expressão:

$$\mathcal{E}[h(\theta; \alpha, \beta)] \sim \frac{1}{2} \log [(\alpha + \beta)^3 / \alpha \beta] - \frac{1}{2} \log 2\pi - \frac{1}{2}$$

donde concluiu que no espaço dos parâmetros, $\{(\alpha, \beta) : \alpha > 0, \beta > 0\}$, as isoentrópicas são dadas aproximadamente, para valores grandes de α e β , por curvas da forma:

$$(\alpha + \beta)^3 / \alpha \beta = \text{constante}$$

tal como se reproduz na figura 4 (a parte a tracejado foi obtida por Lindley por cálculo numérico).

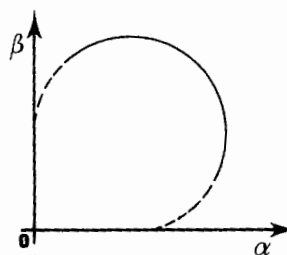


Figura 4

Procurou-se ir um pouco mais longe na identificação das isoentrópicas no espaço dos parâmetros, tentando fazer a sua descrição na região onde as aproximações de Lindley não são legítimas. O cálculo foi feito através de um programa de integração preparado especialmente para o efeito. A figura 5 foi traçada com base em determinações numéricas da entropia efectuadas com esse programa, devendo notar-se — procedimento que em nada altera as intenções da análise — que se trabalhou com entropias negativas ou nega-entropias e que se tomaram logaritmos decimais em vez de logaritmos neperianos. As isoentrópicas são simétricas em relação à recta $\alpha=\beta$ e têm forma ovóide. Além disso, as tangentes às isoentrópicas nos pontos da recta $\alpha=\beta$ são perpendiculares a esta recta (v. secção 4), o que parece estar em oposição à parte a tracejado na determinação de Lindley; para α e β grandes as curvas da figura 5 têm andamento semelhante ao da figura 4.

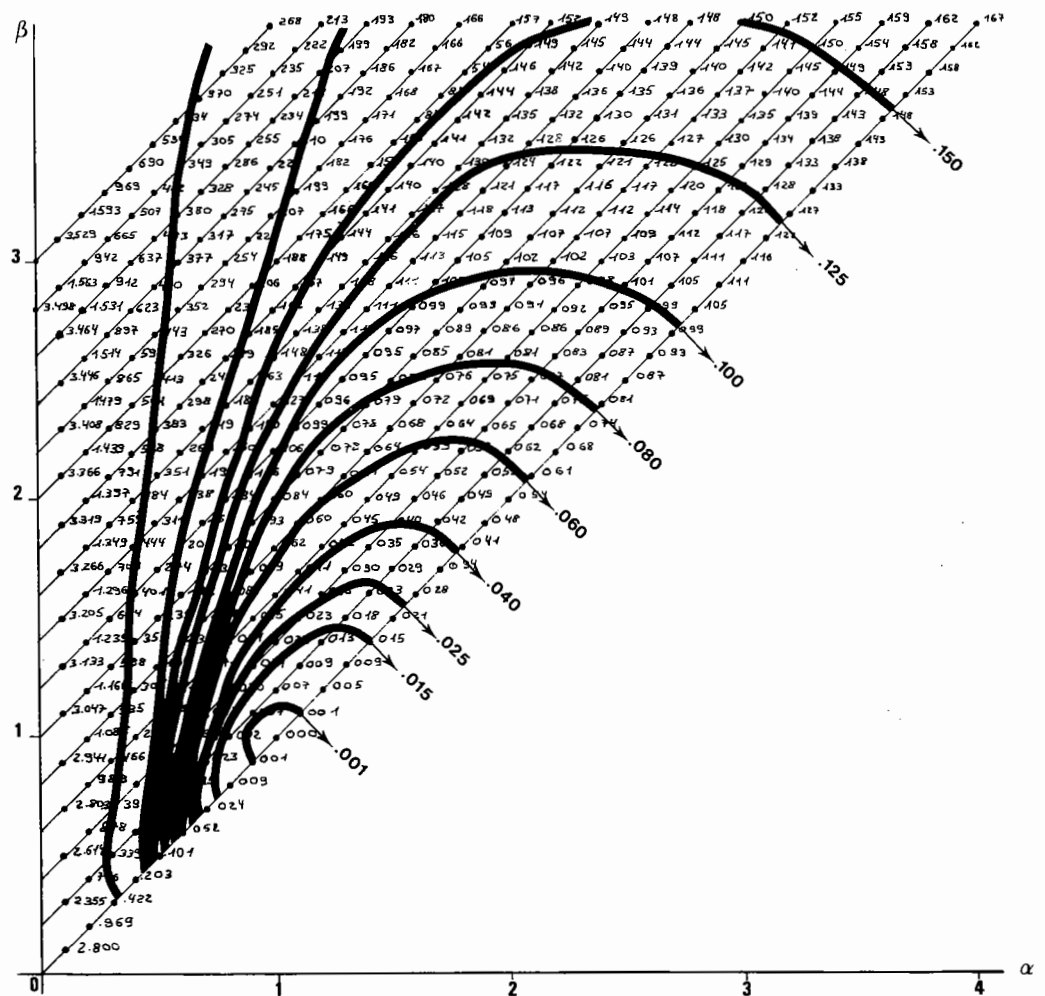


Figura 5

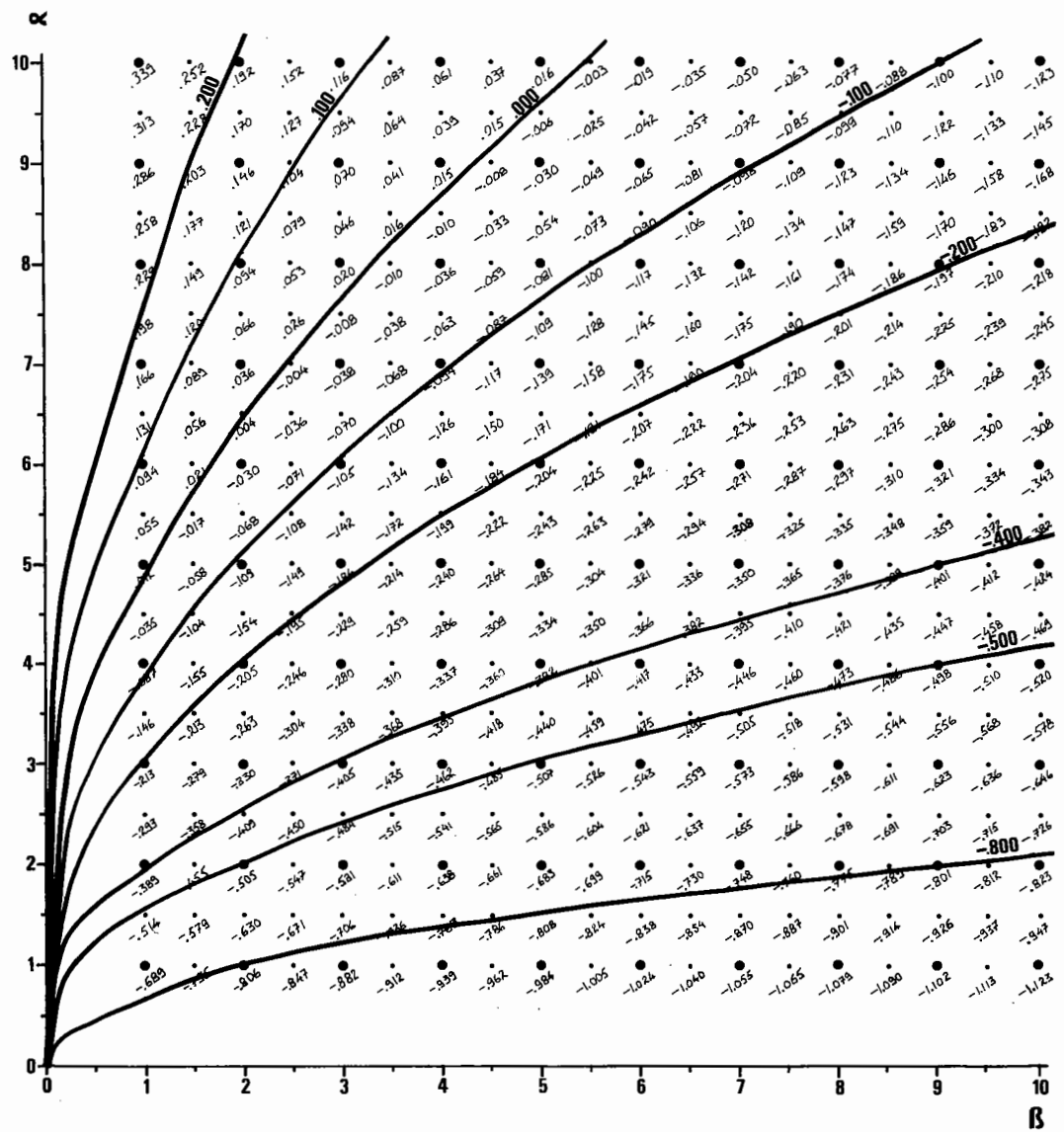


Figura 6

A figura 5 permite já identificar uma diferença em relação ao caso da normal. Comparando as figuras 1 e 5, verifica-se que, no caso da binomial, alguns resultados experimentais podem fazer passar de um ponto (α_0, β_0) para um ponto (α_1, β_1) com maior entropia. No entanto, como (3.2), elucida, a entropia, em média, nunca pode aumentar.

Outro caso de interesse diz respeito à família conjugada da distribuição de Poisson:

$$(3.9) \quad \mathcal{F}_p = \{f(x|\theta) = e^{-\theta} \theta^x / x! : \theta > 0\}, \mathcal{X} = \{0, 1, 2, \dots\}$$

que se sabe ser a gama:

$$(3.10) \quad \mathcal{H}_p = \{h(\theta; \alpha, \beta) = \alpha^\beta e^{-\alpha\theta} \theta^{\beta-1} / \Gamma(\beta) : \alpha > 0, \beta > 0\}, \Theta = (0, +\infty)$$

Com efeito, se a distribuição *a priori* é $h(\theta) \equiv h(\theta; \alpha_0, \beta_0)$ e se a experiência consiste na observação de uma amostra casual de dimensão N , $X = (X_1, X_2, \dots, X_N)$ de uma população Poisson, sabe-se que a distribuição *a posteriori* é ainda uma gama, $h(\theta|x) \equiv h(\theta; \alpha_1, \beta_1)$, com parâmetros:

$$(3.11) \quad \alpha_1 = \alpha_0 + N, \beta_1 = \beta_0 + \sum x_i$$

A distribuição gama tem entropia:

$$(3.12) \quad \mathcal{E}[h(\theta; \alpha, \beta)] = \log[\alpha / \Gamma(\beta)] - \beta + (\beta - 1) \Psi(\beta)$$

onde $\Psi(u)$ é mais uma vez a derivada logarítmica da função gama ou função digama.

Adaptando o programa de integração já referido, foi possível obter as isoentrópicas no espaço dos parâmetros da distribuição gama. O resultado apresenta-se na figura 6, devendo notar-se, mais uma vez, que se trabalhou com nega-entropias e com logaritmos decimais; além disso, as nega-entropias aparecem na figura adicionadas do valor $-0,255$, o que, como é evidente, em nada altera a estrutura das isoentrópicas. Como se mostra na secção 4, tem-se que, para todas as isoentrópicas, $da/d\beta \rightarrow \infty$, quando $\beta \rightarrow 0$, e $da/d\beta \rightarrow 0$, quando $\beta \rightarrow \infty$.

4 — Complementos sobre isoentrópicas

Considere-se uma família de distribuições com suporte em Θ :

$$\mathcal{H} = \{h(\theta; \lambda) : \lambda \in \Lambda\}$$

onde $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$. Numa dada isoentrópica:

$$(4.1) \quad \int_{\Theta} h(\theta; \lambda) \cdot \log h(\theta; \lambda) d\theta = \int_{\Theta} h(\theta; \lambda + \delta\lambda) \cdot \log h(\theta; \lambda + \delta\lambda) d\theta = C$$

onde C é uma constante.

Em face do desenvolvimento em série:

$$(4.2) \quad h(\theta; \lambda + \delta\lambda) = h(\theta; \lambda) + \sum_i \delta\lambda_i [\partial h(\theta; \lambda) / \partial \lambda_i] + \dots$$

fazendo $\partial h(\theta; \lambda) / \partial \lambda_i = \partial_i h(\theta; \lambda)$, tem-se de (4.1).

$$\begin{aligned} & \int_0^1 h(\theta; \lambda) \{ \log [h(\theta; \lambda) + \sum_i \delta\lambda_i \partial_i h(\theta; \lambda)] - \log h(\theta; \lambda) \} d\theta + \\ & + \int_0^1 \sum_i \delta\lambda_i \partial_i h(\theta; \lambda) \cdot \log [h(\theta; \lambda) + \sum_i \delta\lambda_i \partial_i h(\theta; \lambda)] d\theta = 0 \end{aligned}$$

ou ainda:

$$(4.3) \quad \begin{aligned} & \int_0^1 h(\theta; \lambda) \cdot \log \{ 1 + [\sum_i \delta\lambda_i \partial_i h(\theta; \lambda) / h(\theta; \lambda)] \} d\theta + \\ & + \int_0^1 \sum_i \delta\lambda_i \partial_i h(\theta; \lambda) \cdot \log [h(\theta; \lambda) + \sum_i \delta\lambda_i \partial_i h(\theta; \lambda)] d\theta = 0 \end{aligned}$$

Recordando que $\log(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \dots$ ($-1 < z \leq 1$), desprezando termos de 2.^a ordem ou de ordem superior, pode escrever-se:

$$\begin{aligned} & \int_0^1 h(\theta; \lambda) [\sum_i \delta\lambda_i \partial_i h(\theta; \lambda) / h(\theta; \lambda)] d\theta + \\ & + \int_0^1 \sum_i \delta\lambda_i \partial_i h(\theta; \lambda) \cdot \log h(\theta; \lambda) \{ 1 + [\sum_i \delta\lambda_i \partial_i h(\theta; \lambda) / h(\theta; \lambda)] \} d\theta = 0 \end{aligned}$$

e, finalmente, depois de simplificar:

$$(4.4) \quad \sum_i \delta\lambda_i [\int_0^1 \partial_i h(\theta; \lambda) d\theta + \int_0^1 \partial_i h(\theta; \lambda) \cdot \log h(\theta; \lambda) d\theta] = 0$$

A expressão (4.4) mostra a relação que existe entre os acréscimos paramétricos, $\delta\lambda_i$, quando se caminha ao longo de uma isoentrópica.

Na aplicação à família beta tem-se (reparametrizando, para facilitar a escrita):

$$\begin{aligned} h(\theta; \alpha, \beta) &= B^* \theta^\alpha (1-\theta)^\beta, \quad B^* = 1/B(\alpha+1, \beta+1), \\ \partial_\alpha h(\theta; \alpha, \beta) &= B^* \theta^\alpha (1-\theta)^\beta \log \theta + (\partial B^* / \partial \alpha) \theta^\alpha (1-\theta)^\beta, \\ \partial_\beta h(\theta; \alpha, \beta) &= B^* \theta^\alpha (1-\theta)^\beta \log(1-\theta) + (\partial B^* / \partial \beta) \theta^\alpha (1-\theta)^\beta \end{aligned}$$

De (4.4) sai, com todos os integrais no intervalo (0, 1):

$$0 = \delta\alpha \{ \int \partial_\alpha h d\theta + \int \partial_\alpha h \cdot \log h d\theta \} + \delta\beta \{ \int \partial_\beta h d\theta + \int \partial_\beta h \cdot \log h d\theta \}$$

Finalmente, no caso particular $\alpha = \beta$, isto é, para os pontos da diagonal do 1.^o quadrante, tem-se $-\delta\beta/\delta\alpha = 1$: as tangentes às isoentrópicas nesses pontos são perpendiculares à referida diagonal. Para chegar a esta conclusão basta notar a simetria da função gama e a invariância que se verifica quando θ é substituído por $(1-\theta)$.

No caso da família normal, $h(\theta; \mu, \sigma) = (1/\sigma\sqrt{2\pi}) \exp\{-(\theta - \mu)^2/2\sigma^2\}$ onde $-\infty < \theta < +\infty$, $-\infty < \mu < +\infty$, $\sigma > 0$, vem:

$$\partial_\mu h = h [(\theta - \mu) / \sigma^2], \quad \partial_\sigma h = (-1/\sigma) h [(\theta - \mu)^2 / \sigma^3]$$

Assim, por (4.4):

$$0 = \delta \mu \left[\int \partial_{\mu} h (1 + \log h) d\theta \right] + \delta \sigma \left[\int \partial_{\sigma} h (1 + \log h) d\theta \right]$$

donde se conclui facilmente que o coeficiente de $\delta \mu$ é igual a zero, por se tratar do integral de uma função ímpar. Logo, $\delta \sigma / \delta \mu = 0$ e as isoentrópicas são rectas paralelas ao eixo dos $\mu \mu$, como, aliás, se tinha concluído por cálculo directo.

Na aplicação à família gama tem-se:

$$h(\theta; \alpha, \beta) = \alpha^{\beta} e^{-\alpha \theta} \theta^{\beta-1} / \Gamma(\beta),$$

$$\partial_{\alpha} h(\theta; \alpha, \beta) = [\beta \alpha^{\beta-1} e^{-\alpha \theta} \theta^{\beta-1} / \Gamma(\beta)] - [\alpha^{\beta} e^{-\alpha \theta} \theta^{\beta} / \Gamma(\beta)],$$

$$\begin{aligned} \partial_{\beta} h(\theta; \alpha, \beta) &= [\alpha^{\beta} e^{-\alpha \theta} \theta^{\beta-1} \log \alpha / \Gamma(\beta)] + [\alpha^{\beta} e^{-\alpha \theta} \theta^{\beta-1} \log \theta / \Gamma(\beta)] - \\ &\quad - [\alpha^{\beta} e^{-\alpha \theta} \theta^{\beta-1} \frac{1}{\Gamma(\beta)^2} \frac{d \Gamma(\beta)}{d \beta}] \end{aligned}$$

donde, por (4.4):

$$-\frac{\delta \alpha}{\delta \beta} = \frac{\int [\log \alpha + \log \theta - \Psi(\beta)] h(1 + \log h) d\theta}{\int [(\beta / \alpha) - \theta] h(1 + \log h) d\theta}$$

Atendendo-se à segunda expressão de (3.8), conclui-se facilmente que $\delta \alpha / \delta \beta \rightarrow 0$, quando $\beta \rightarrow \infty$; por outro lado, como $\Psi(\beta) \rightarrow -\infty$, quando $\beta \rightarrow 0$, tem-se $-\delta \alpha / \delta \beta \rightarrow \infty$, quando $\beta \rightarrow 0$.

Fica assim algo confirmada a forma das isoentrópicas da figura 6.

BIBLIOGRAFIA

- BARNETT, V. — *Comparative Statistical Inference*, Wiley, London, 1973.
 BERGER, James O. — *Statistical Decision Theory*, Springer-Verlag, New York, 1980.
 DE GROOT, Morris H. — *Optimal Statistical Decisions*, McGraw-Hill, New York, 1970.
 LINDLEY, D. V. — «On a measure of the information provided by an experiment», *The Annals of Mathematical Statistics*, vol 27, no. 4, 1956.
 PUGACHEV, V. S. — *Introducción a la Teoría de las Probabilidades*, Editorial MIR, Moscovo, 1973.
 VENTZEL, H. — *Théorie des probabilités*, Éditions MIR, Moscovo, 1973.

